



# A hybrid (iron–fat–water) phantom for liver iron overload quantification in the presence of contaminating fat using magnetic resonance imaging

Nazanin Mobini<sup>1,2</sup> · Malakeh Malekzadeh<sup>2,3</sup> · Hamidreza Haghghatkhah<sup>4</sup> · Hamidreza Saligheh Rad<sup>1,2</sup> 

Received: 25 June 2019 / Revised: 29 October 2019 / Accepted: 30 October 2019 / Published online: 16 November 2019  
© European Society for Magnetic Resonance in Medicine and Biology (ESMRMB) 2019

## Abstract

**Objective** Assessment of iron content in the liver is crucial for diagnosis/treatment of iron-overload diseases. Nonetheless,  $T_2^*$ -based methods become challenging when fat and iron are simultaneously present. This study proposes a phantom design concomitantly containing various concentrations of iron and fat suitable for devising accurate simultaneous  $T_2^*$  and fat quantification technique.

**Materials and methods** A 46-vial iron–fat–water phantom with various iron concentrations covering clinically relevant  $T_2^*$  relaxation time values, from healthy to severely overloaded liver and wide fat percentages ranges from 0 to 100% was prepared. The phantom was constructed using insoluble iron (II, III) oxide powder containing microscale particles.  $T_2^*$ -weighted imaging using multi-gradient-echo (mGRE) sequence, and chemical shift imaging spin-echo (CSI-SE) Magnetic Resonance Spectroscopy (MRS) data were considered for the analysis.  $T_2^*$  relaxation times and fat fractions were extracted from the MR signals to explore the effects of fat and iron overload.

**Results** Size distribution of iron oxide particles for Magnetite fits with a lognormal function with a mean size of about 1.17  $\mu\text{m}$ . Comparison of FF color maps, estimated from bi- and mono-exponential model indicated that single- $T_2^*$  fitting model resulted in lower NRMSD. Therefore,  $T_2^*$  values from the mono-exponential signal equation were used and expressed the relationship between relaxation time value across all iron (Fe) and fat concentration as  $\text{Fe} = -28.02 + \frac{302.84}{T_2^*} - 0.045 \text{FF}$ , with R-squared = 0.89.

**Discussion** The proposed phantom design with microsphere iron particles closely simulated the single- $T_2^*$  behavior of fatty iron-overloaded liver in vivo.

**Keywords** Iron–fat–water phantom · Magnetic resonance imaging · Iron overload · Fatty liver

## Introduction

Iron is an essential mineral found in all body cells [1]. Although iron is excreted through a variety of ways including bleeding, sweating, etc., there is no mechanism in the body to adjust excess iron removal [2, 3]. Hence, in some conditions such as hereditary hemochromatosis (genetic mutations), or hemosiderosis (repeated blood transfusions), iron overload may occur which is more damaging than iron deficiency [4]. Excess iron is mainly stored in the liver [3, 5] and might be concurrent with increase in hepatic fat percentage (steatosis) that exacerbates the disease condition [6–8]. Unfortunately, iron deposition has no early specific symptoms and if missed, it can cause organs' dysfunction.

The gold standard method for estimating liver iron concentration (LIC) and steatosis level in vivo is liver

✉ Hamidreza Saligheh Rad  
h-salighehrad@tums.ac.ir

<sup>1</sup> Department of Medical Physics and Biomedical Engineering, School of Medicine, Tehran University of Medical Sciences (TUMS), Tehran, Iran

<sup>2</sup> Quantitative MR Imaging and Spectroscopy Group (QMISG), Research Center for Cellular and Molecular Imaging, Institute for Advanced Medical Technologies, Imam Khomeini Hospital Complex, Tehran University of Medical Sciences, Keshavarz Blvd, 1419733141 Tehran, Iran

<sup>3</sup> Medical Physics Department, School of Medicine, Iran University of Medical Sciences (IUMS), Tehran, Iran

<sup>4</sup> Department of Radiology, Shohada-e Tajrish Hospital, Shahid Beheshti University of Medical Sciences (SBMU), Tehran, Iran

biopsy. LIC results reported from biopsy are expressed as milligram of Fe to the gram of dry-weight-tissue (mg Fe/gdw). Accordingly, iron overload can be categorized into four grades: normal (< 3 mg Fe/gdw), mild (3–7 mg Fe/gdw), moderate (7–15 mg Fe/gdw) and severe (> 15 mg Fe/gdw) [4]. Similarly, hepatic steatosis based on the rate of affected cells or fat percentage can be categorized into grade 0 (< 5%), grade 1, (5–33%), grade 2 (34–66%) and grade 3 (> 67%) [7, 9]. The invasive and painful nature of the biopsy procedure, as well as the proneness to sampling errors due to the uneven distribution of iron and fat, makes it inappropriate for repeated measurements [10, 11]. Hence, there is a need for alternative non-invasive evaluation methods to biopsy.

In this context, magnetic resonance imaging (MRI) as a non-ionizing imaging modality and a non-invasive measurement technique with sensitivity to both iron and fat deposition can be employed. Iron particles distort uniformity of the main magnetic field, which results in dephasing of the signal from ensembles of spins and decreases the received gradient recalled echoes (GRE) MR signal [12–14]. Additionally, fat spectrum is the superposition of multiple peaks with different frequencies. Concurrent existence of fat in the hepatocytes imposes oscillation to the exponential decay behaviour of the signal and further reduces the total  $T_2^*$ . Therefore, GRE signal has the potential to detect iron and fat simultaneously which could help in choosing the appropriate therapy and also treatment follow-ups [15].

Nowadays, various techniques of MRI are used to assess liver iron concentration. However, non-similar results in different imaging centers, inability to measure iron in severe conditions, errors in concomitant presence of fat and iron and limitation in providing biopsy references to evaluate accuracy of new methods restrict the clinical utility of MRI. In order to deal with these issues, phantoms must be designed to simulate concomitant liver fat and liver iron conditions.

Most previous attempts in phantom construction focused on quantifying fat or iron separately, such as the study of iron overload with a homogeneous phantom containing series of different manganese chloride ( $\text{MnCl}_2$ ) concentrations by Clark et al. [16], or a study of fat–water phantom with variation in fat content from 0 to 100% by Bernard et al. [17]. The design of the phantom was improved by Hines et al. through considering simultaneous presence of fat (0–100%) and iron (Feridex; 0–50  $\mu\text{g Fe/mL}$ ). The results of this study revealed that nanoscale iron particles could not significantly affect the magnetic field of fat droplets. Therefore,  $T_2^*$  of water and fat were diverse and different decay times or dual- $T_2^*$  behaviour were observed [18]. Recent in vivo MR experiments on fatty iron-overloaded tissue represented similar signal decay times for water and fat, implying single- $T_2^*$  or mono-exponential signal decay behaviour [6, 9, 19].

In this study, we aimed to further improve the design of iron–fat–water from a magnetic resonance signal perspective by exploiting iron particles with sizes at least a full order of magnitude larger than Feridex. Contrary to the earlier homogeneous phantom composition that included iron oxide (Magnetite) aqueous suspension or SPIO, magnetite micro-particles have powder insoluble physical form that will bring about a non-homogenous compound. Considering the fact that the excess accumulated iron particles in the liver are in the form of ferritin clusters and their sizes are in the order of micrometres rather than nanometres, using microsphere iron oxide particles might better simulate liver iron overload [19]. Also, similar to the studies performed by Bernard et al. [17] and Hines et al. [18], the effect of various fat concentrations was considered in our proposed phantom. The phantom with a better mimic of different liver disease conditions can be applied to provide a method for accurate  $T_2^*$  quantification studies.

## Materials and methods

### Measurement of iron particle size

Initially, morphological properties of the iron oxide powder were studied. Measurement of the magnetite particles was accomplished using a Field Emission Scanning Electron Microscope (Fe-SEM; HITACHI S4160 with 5 nm resolution). Images were captured with this high-resolution microscope as follows: first, in order to sample preparation for electron microscopy, the dry iron oxide powder was coated with gold to provide a conductive surface. Then, the prepared sample was scanned by a narrow beam of electrons and an image was formed on the monitor that expressed the morphology of the iron oxide in surface layers [20].

Relative particle size is usually determined using image analysis techniques that completely rely on the quality of SEM pictures. Because there were many overlapping neighbours that made us exploit manual evaluation, a few hundred individual particles were selected and evaluation was performed by measuring the particle sizes to obtain the distribution and to estimate the geometric average value.

### Phantom design

A non-homogenous iron–fat–water phantom, consisting of forty-six 50-mL vials of water solution approximately similar to that proposed by Hines et al. [18] was made as follows: 0.8% mass/volume concentration agar (in order to obtain stable emulsions in the form of gel, not solid), 0.5 mM cupric sulphate (to reduce  $T_1$  value of the solution to approximately 800 ms at 3.0 T to be more similar to biological values of  $T_1$ ), 43 mM sodium chloride, 3.75 mM sodium

azide (antimicrobial agent) and 43 mM sodium dodecyl sulphate (surfactant or emulsifier) all dissolved in distilled deionized water. Water solution was mixed with appropriate proportions of magnetite and peanut oil. Magnetite is a kind of iron oxide with the chemical formula  $\text{Fe}_3\text{O}_4$ , has 72.36% iron by mass. Thus, final iron concentrations were 0, 14.5, 43.4, 72.4, 101.3, 144.7, 217.1, 361.8 and 723.6  $\mu\text{g}$  Fe/mL selected to cover clinically relevant  $T_2^*$  relaxation time values, from healthy (about 20 ms) to severely overloaded liver (less than a few milliseconds). Peanut oil concentrations were varied to mimic fat percentages of 0%, 5%, 15%, 30%, 50% and 100%. The mixture was heated to boil for 30 s, shaken carefully and cooled down to form a soft gel. These procedures allowed non-soluble  $\text{Fe}_3\text{O}_4$  particles to suspend, and other compounds to mix in the solution. It was important that the final solution was well-mixed so that the boundary between oil and water disappeared similar to fatty liver, where fat compounds are unobservable. This gel phantom is stable at room temperature and does not require special maintenance.

Vials containing many observable deposited iron particles in the bottom of the containers or those with visible borders between water and fat were excluded from our quantitative study.

## Image acquisition

The phantom was imaged at a 3.0 T scanner (MAGNETOM Prisma, Siemens, Erlangen, Germany), using an 18-channel body array coil. First series of magnitude coronal  $T_2^*$  images were acquired using a 2D multi gradient recalled echoes (mGRE) pulse sequence, with the imaging parameters as follows: TR=400 ms with a  $20^\circ$  flip angle that was deliberately chosen non-optimal to decrease bias,  $\text{TE}_1 = 1.71$  ms,  $\Delta\text{TE} = 0.74$  ms, sampling bandwidth = 1930 Hz per pixel, number of average (NEX) = 12 (to minimize noise), voxel size =  $3.9 \times 3.9 \times 10.0$  mm<sup>3</sup>, FOV = 370 mm, matrix =  $66 \times 96$ , TA = 11 min 02 s and the number of slices = 10. Then, the second and thirds series of images were acquired using the same protocol, but with the first TEs equal to 2.05 and 10.49 ms, respectively. Through combination of these separate acquisitions, 30 non-equally spaced TEs from 1.71 to 14.2 ms were provided with denser echo time spaces.

Additionally, a vial with 100% fat content was employed for coronal 1H magnetic resonance spectroscopy (1H-MRS), imaged with chemical shift imaging spin echo (CSI\_SE) sequence using with TE/TR = 30/1200 ms (long TR to minimize the  $T_1$  effect), voxel size =  $5.8 \times 4.6 \times 40.0$  mm<sup>3</sup>, transverse magnetization (TM) = 0 ms, TA = 5 min 58 s, BW = 1000 Hz, NEX = 10, frequency of H = 123.19 MHz, without any fat or water suppression.

## MR data analysis

The GRE signal containing fat and iron can be modelled as a function of  $TE$  as follows [6]:

$$S(\text{TE}_n) = \left( W \cdot \exp\left(\frac{-\text{TE}_n}{T_{2W}^*}\right) + F \cdot \exp\left(\frac{-\text{TE}_n}{T_{2F}^*}\right) \sum_{m=1}^M \alpha_M \cdot \exp(j2\pi f_m \text{TE}_n) \right) \exp(j2\pi f_B \text{TE}_n) + \eta, \quad (1)$$

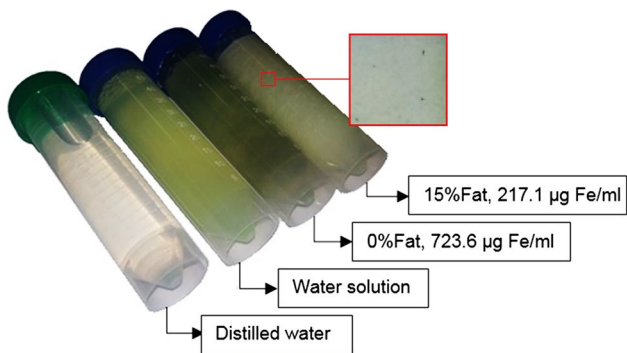
where  $S(\text{TE}_n)$  is the complex signal acquired at  $\text{TE}_n$ ,  $W$  and  $F$  are the water and fat amplitudes at  $\text{TE} = 0$  (fat fraction is  $\text{FF} = \frac{100F}{F+W}$ ),  $T_{2W}^*$  is the relaxation time of water,  $T_{2F}^*$  is the relaxation time of fat,  $f_B$  represents the field inhomogeneity and  $\eta$  is an additive complex Gaussian noise (zero mean and standard deviation of  $\sigma_0$ ). Moreover, the terms inside the sigma describe fat characteristics,  $\alpha_M$  ( $\sum_{m=1}^M \alpha_M = 1$ ) and  $f_M$ , are the relative amplitude and frequency location of  $M$ -th peak visible in the spectrum, respectively. Analyzing shape of the acquired MRS dataset with the help of jMRUI (AMARES) software package enabled us to pre-calibrate a 6-peak spectral model of the fat that needed to be performed just once. Placing the spectrum in the signal model simplified the calculation of estimating other unknown parameters.

By employing an in-house algorithm developed in MATLAB Release 2016b (The MathWorks Inc.),  $T_2^*$ -weighted images were analysed to calculate pixel-by-pixel color maps. In order to remove background pixels and to calculate the average parameter values of the vials, 45 regions of interest (ROIs), each surrounding one of the vials, were drawn on an image slice. Calculations were performed with both dual- and single- $T_2^*$  equation to compare accuracy, where distinct and afterward common relaxation values were assumed for water and fat, respectively. Magnitude form of the described signal model was used for analysis which is more common in the clinic and keeps noise-induced bias away in FF estimation. The initial amount of  $W$  was chosen to be the highest signal intensity at each pixel. Moreover, to set the best initial  $T_2^*$  value, two primary fitting procedures were performed: first on signal samples where the water and fat signal seemed to be in phase (i.e.  $\text{TE} = [2.45, 4.67, 6.89, 9.11]$ ), then on samples that seemed to be out of phase (i.e.  $\text{TE} = [3.53, 5.75, 7.97, 10.19]$ ). Then, the average estimated  $T_2^*$  was used as the best initial choice for the final fitting result. Implementation of the fitting with nonlinear-least square solver using Trust-Region-Reflective algorithm resulted in 2 distinct  $T_2^*$  maps for the bi- and a  $T_2^*$  map for the mono-exponential signal model, as well as a FF map for each of the equations. In addition, a multiple regression was conducted to investigate the  $T_2^*$  relaxation behaviour across all iron and fat contents at 3.0 T.

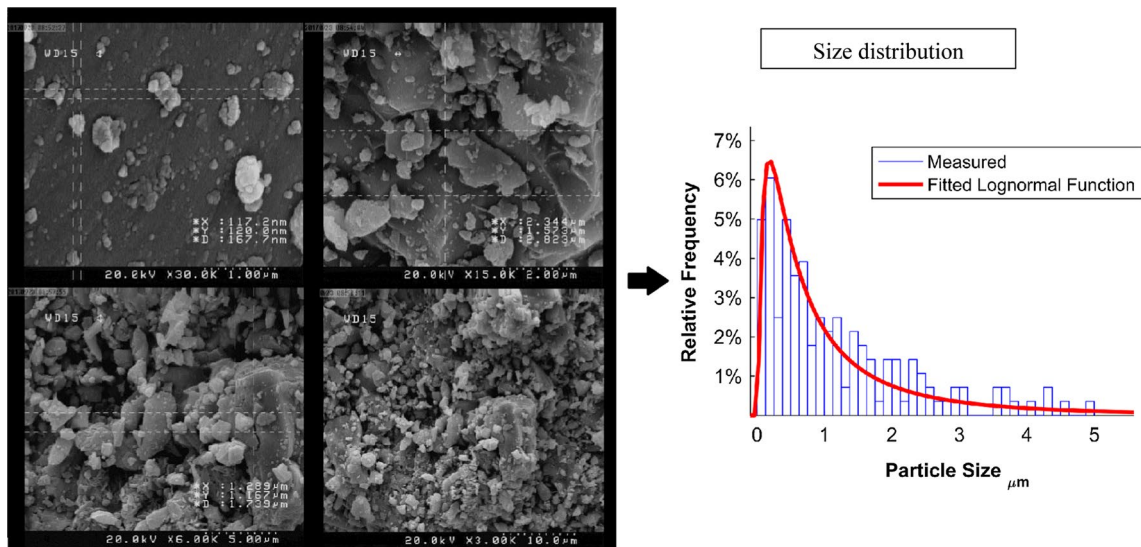
To compare forecasting errors of bi- and mono-exponential models for the multi-GRE dataset, normalized root-mean-square deviation (NRMSD) was used. NRMSD as a measure of accuracy represented magnitudes of the errors or residuals between the true FF and the estimated FF maps that resulted from each of the two models.

## Results

A couple of vials with different contents are shown in Fig. 1 which illustrates the phantom preparation has been a successful process. Like the real tissue, no boundary is observed between water and fat. Moreover, the gel-based solution has



**Fig. 1** Four selected (among a total of 46) vials. The resultant fat-water-iron solution is heterogeneous and the suspended iron particles are visible. Moreover, there's no boundary between water and fat (same as tissue)



**Fig. 2** FE-SEM micrographs (in four different scales) and particle size distribution of iron oxide powder. The particle size distribution histogram of the sample shows a lognormal pattern with average size of 1.17  $\mu\text{m}$

provided a substrate for suspending iron particles that prevented them from settling out.

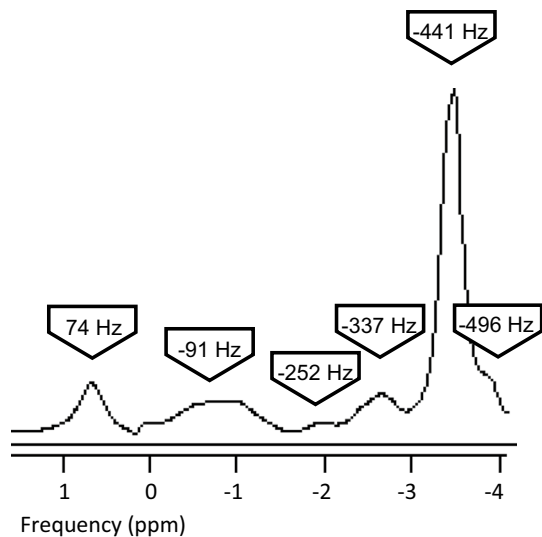
Figure 2 displays SEM images of the used magnetite sample and particle size distributions determined from the images. Manual analysis estimated that the average iron oxide particle size of the sample was 1.17  $\mu\text{m}$ , 34.67% of the sample contained particles smaller than 500 nm and sizes of the rest of the particles fell within the range of 0.5–5  $\mu\text{m}$ .

Figure 3 is a representative example of the fat spectrum which was used to measure relative amplitude and frequency location of each peak. The characteristic information of the peaks was placed in the signal equation to decrease the number of unknowns.

Figure 4 shows pixel-by-pixel  $T_2^*$  and FF colour maps, estimated from bi-exponential model in the first row and mono-exponential model in the next. Vials were arranged in the order in which iron concentration increases from bottom to up and fat percentage increases from right to left. MRI experiments on the phantom proved that  $T_2^*$  relaxation time decreases in the presence of suspended microsphere iron oxide particles and in the presence of fat with zero iron concentration. The localizer of acquisition had adjusted such that the bottom of vials or the effect of iron particles deposited at the bottom of vials was not included in the slices; therefore, the vial with 0% fat and 101.3  $\frac{\mu\text{gFe}}{\text{mL}}$  iron concentration with non-dispersed iron particles was excluded from subsequent analysis.

In Fig. 5, NRMSD or measure of the differences between true FF values and the estimated FF by the dual- and single- $T_2^*$  signal models are depicted. Estimated FF maps show that dual- $T_2^*$  fitting model resulted in poor outcomes and





**Fig. 3** Illustrative multi-peak spectrum acquired at 3.0 T in peanut oil that was used as fat in the phantom. The chemical shift frequency of each peak is labelled in Hz relative to water

high standard deviations, particularly in vials with high iron contents.

Accordingly, results derived from the mono-exponential model that were shown in the second row of Fig. 4 are summarized in Fig. 6, where mean estimated  $T_2^*$  relaxation time (average of relaxations within each of the ROIs) and mean estimated PDFF in each vial are plotted against the relative iron concentration and true fat percentage, respectively. It was observed that in the last three high iron concentrations (i.e.  $\geq 217.1 \frac{\mu\text{gFe}}{\text{mL}}$ ), the estimated relaxation values reached saturation. Moreover, in the PDFF graph, the deviation between the estimated fat percentage and the true value increases by the increment in iron content which is completely visible in the  $723.6 \frac{\mu\text{gFe}}{\text{mL}}$  iron concentration. Therefore, the results from the three highest iron contents were eliminated and the remaining  $T_2^*$  values following a multiple regression models the relationship between relaxation time value over all iron (Fe) and fat concentration as  $\text{Fe} = -28.02 + \frac{302.84}{T_2^*} - 0.045\text{FF}$ , with R-squared = 0.89. Coefficients were determined with 95% confidence bounds.

## Discussion

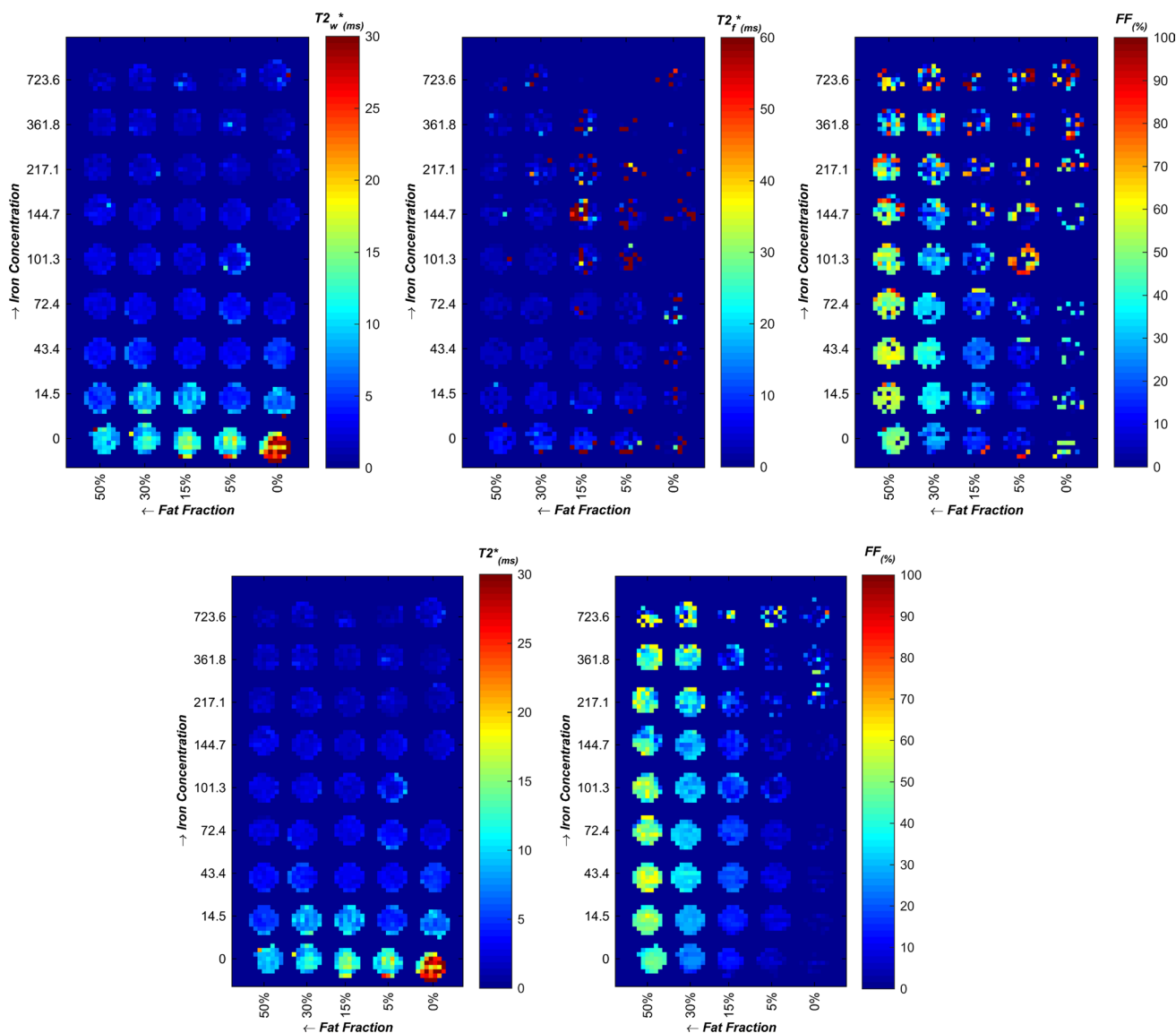
In this study, a gel phantom containing a mixture of water solution and fat and with a broad range of iron was prepared. To fix the issues with previously proposed phantoms [18], microsphere iron oxide sample with larger particles than SPIO was used to get slightly closer behaviour to the

actual liver condition. MRI experiments on this phantom indicated that  $T_2^*$  relaxation time decreases in the presence of suspended microsphere iron oxide particles.

As demonstrated by the results of the two fitting models, mono-exponential equation results in lower error than the dual- $T_2^*$  fitting, which was inline with the recent research; the decay rates of water and fat signals in diseased human liver and animal models with concomitant hepatic steatosis and iron overload are very similar [6, 19]. Contrary to the earlier homogeneous phantom that included iron oxide aqueous suspension, it seems that insoluble microsphere iron particles affected both water and fat; therefore, signal from each voxel was better fitted by a mono-exponential function. In other words, our proposed phantom with identical  $T_{2^*W}$  and  $T_{2^*F}$  behaviour mimics liver better [6].

The amounts of iron in phantom are similar to wet tissue, and since the fraction of water in soft tissues is around 70–85%, they must be converted to dry-weight by a factor of 5.5 in order to be comparable with the in vivo results [21, 22]. In the present phantom, we have tried to simulate healthy-to-severe iron overload conditions from a magnetic resonance signal perspective or clinically relevant relaxation times through an extended iron content ( $0\text{--}723.6 \mu\text{g Fe/mL} \sim 4 \text{ mgFe/gdw}$ ), though these concentrations can just cover the physiological liver iron range up to the mild level. In other words, excess iron particles accumulated in the liver have less impact on the decay rate of water molecules than what we observed in our phantom and accordingly, the phantom reached to less than a few milliseconds with nearly mild hepatic iron amount. This dissimilarity between the in vivo intracellular decay time and phantom may be due to the microscopic characteristics of the aqueous fat/iron solution with the uncontrolled size of fat droplets and iron clusters. Although it is not of interest to reach precise physiologic iron-overload amounts, it is important to cover  $T_2^*$  value created by excess iron in different grades of liver disease. The generated regression curve which has modeled the iron content as a function of estimated relaxation times and fat percentage could be further improved by calibrating with the results obtained by patient biopsies to create a calibration curve specified for each imaging protocol intended for direct iron/fat quantification applications in clinics [13].

Furthermore, the results in vials with zero iron concentration indicate that the  $T_2^*$  relaxation values are shortened by fat which is similar to the results reported by Fukuzawa et al. [23]. A recent study by Bydder et al. [24] has pointed out sources of bias in PDFF estimation using magnitude image analysis. It is shown that the  $R_2^*$  difference between water and fat could be replaced with a constant term which is



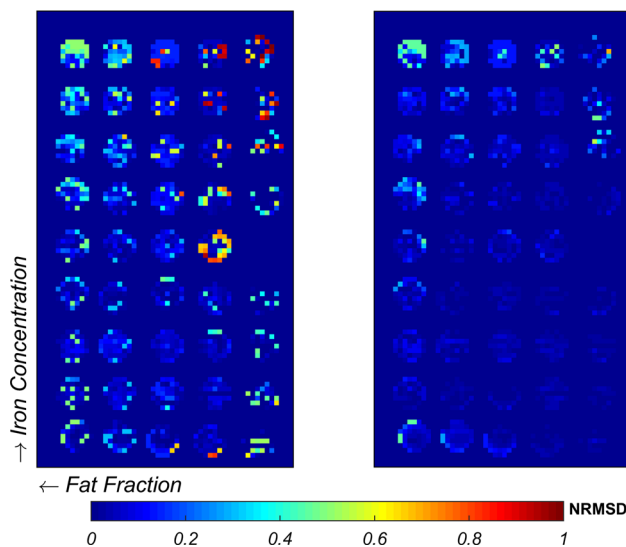
**Fig. 4** Pixel-by-pixel  $T_2^*$  and FF maps, with dual- (first row) and single- $T_2^*$  model (second row). Iron concentration increases from bottom to up and fat percentage increases from right to left. Note that

iron particles did not suspend well in the vial with 0% fat and 101.3  $\frac{\mu\text{grFe}}{\text{mL}}$  iron concentration, so the vial was removed from further analysis

negligible at low to moderate fat percentages and, therefore, the mono-exponential  $T_2^*$  signal model would be sufficient. This small constant also indicates broader linewidth for the liver fat peaks than water, which reduces the total  $T_2^*$  of fat–water mixture. Further studies are needed in order to unravel the exact underlying principle of this observation.

Some limitations of this work should be addressed. First,  $T_2^*$  value for healthy liver with previous studies on 1.5 T MRI is 25–35 ms. In severe iron-overload conditions, signal intensity decreases and the relaxation time value diminishes

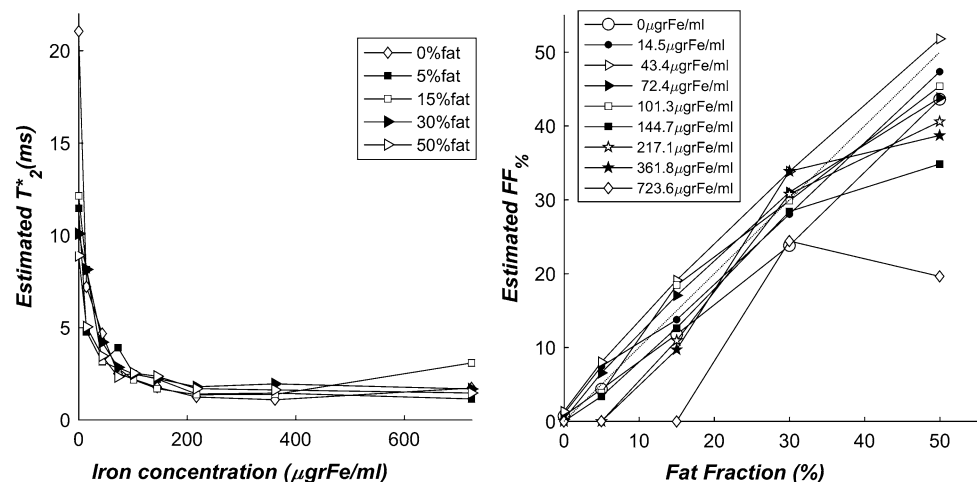
to less than 2 ms and thus it becomes less than 1 ms at 3.0 T scanners [25–27]. The signal loss is reasonable due to equation of the MR signal where  $T_2^*$  is the exponential time constant, and after a period of 3 time constants, only 5% of the initial signal remains. Poor signals could be dominated by noise and complicate quantitative analysis. Therefore, TEs or at least the first TE must be selected shorter than  $T_2^*$  to capture this fast decaying signal. We tried to choose TE as small as possible, but this choice was restricted by the scanner. Also, for the least possible initial echo time, we



**Fig. 5** Normalized RMSE measure of FF estimated with (left) bi- and (right) mono-exponential signal models

had to increase the bandwidth. As a result, signal-to-noise was lost, which was compensated by considering a larger voxel size. Ultimately, we reached  $TE = 1.71$  ms that lead to overestimation in  $T_2^*$  values for vials with high iron content (i.e.  $\geq 217.1 \frac{\mu\text{grFe}}{\text{mL}}$ ) because most of their signal is decayed even at the minimum echo time. To take advantage of high magnetic field beside the sensitivity of  $T_2^*$  to iron, ultrashort echo time imaging could be performed. UTE imaging could decrease TE and also  $\Delta TE$  down to microseconds [28, 29]. Shorter echo time would potentially provide stronger signal and could conquer some SNR problems, especially at high iron concentrations.

**Fig. 6** Results derived from the mono-exponential model inside each ROI: mean  $T_2^*$  values against iron concentrations in varying fat percentages (left), and mean estimated PDFF values against true PDFF in varying iron contents (right)



Second, there were several problems in phantom preparation such as weighing up a small amount of iron powder, due limitation of the scale resolution to 0.1 mg. Therefore, we could not prepare vial contents with less than  $14.5 \mu\text{gr Fe/mL}$  concentration (least possible amount = 1 mg). Moreover, although the non-uniform suspension of Magnetite iron particles in the vials was in line with the studies on explanted liver and heterogeneity of hepatic iron concentration [12], different microscopic iron clustering may result in different  $T_2^*$  decay that might be an additional confounder since it was proved to associate with both the concentration and distribution of iron-oxide particles [14]. Using an aqua sample of microsphere iron with homogenous dispersion in the water solution could reduce the error of phantom and improve the accuracy of regression curve for direct use in clinic.

## Conclusion

In conclusion, an iron–fat–water phantom with broad range of iron was employed in this study that showed single- $T_2^*$  behaviour behaves same as fatty iron-overload liver in vivo. The described phantom could further be used for evaluation of MRI quantification algorithms to improve the profitable use of this technique in clinics, particularly in complex conditions when fat and high iron concentration (with significantly reduced  $T_2^*$ ) simultaneously coexist. Quantification techniques could be further improved by complex fitting to discriminate fat and water signal, performing patient studies and benchmarking the values with biopsies.

**Acknowledgements** The authors would like to thank Hamid Emadi for help in preparing the phantom, Anahita Fathi Kazerooni for her scientific editing and comments on the manuscript. Imaging for this work was performed at National Brain Mapping Laboratory (NBML).

**Author contributions** NM was responsible for study conception and design, data collection, analysis and interpretation of data, and drafting the manuscript; MM helped with the interpretation of data and critical revision; HH designed the study conception and acquisition of data; HS was responsible for protocol/project development of the framework and critical revision.

**Funding** This phantom study has been supported by Tehran University of Medical Sciences & Health Services Grants 27331 and 32965.

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

**Ethical approval** This article does not contain any studies with human participants or animals performed by any of the authors.

## References

- Gupta CP (2014) Role of iron (Fe) in body. *IOSR J Appl Chem (IOSR-JAC)* 7:38–46
- Andrews NC (1999) Disorders of iron metabolism. *N Engl J Med* 341(26):1986–1995
- Brittenham GM, Badman DG (2003) Noninvasive measurement of iron: report of an NIDDK workshop. *Blood* 101(1):15–19
- Batts KP (2007) Iron overload syndromes and the liver. *Mod Pathol* 20(1s):S31
- Yokoo T, Yuan Q, Sénégas J, Wiethoff AJ, Pedrosa I (2015) Quantitative R2\* MRI of the liver with Rician noise models for evaluation of hepatic iron overload: Simulation, phantom, and early clinical experience. *J Magn Reson Imaging* 42(6):1544–1559
- Hornig DE, Hernando D, Reeder SB (2017) Quantification of liver fat in the presence of iron overload. *J Magn Reson Imaging* 45(2):428–439
- Yokoo T, Browning JD (2014) Fat and iron quantification in the liver: past, present, and future. *Top Magn Reson Imaging* 23(2):73–94
- Kühn J-P, Meffert P, Heske C, Kromrey M-L, Schmidt CO, Mensel B, Völzke H, Lerch MM, Hernando D, Mayerle J (2017) Prevalence of fatty liver disease and hepatic iron overload in a Northeastern German population by using quantitative MR imaging. *Radiology* 28(3):706–716
- Tipirneni-Sajja A, Krafft AJ, Loeffler RB, Song R, Bahrami A, Hankins JS, Hillenbrand CM (2019) Autoregressive moving average modeling for hepatic iron quantification in the presence of fat. *J Magn Reson Imaging* 50:1620–1632. <https://doi.org/10.1002/jmri.26682>
- Lidbury JA (2017) Getting the most out of liver biopsy. *Vet Clin Small Animal Pract* 47(3):569–583
- Ratziu V, Charlotte F, Heurtier A, Gombert S, Giral P, Bruckert E, Grimaldi A, Capron F, Poynard T, Group LS (2005) Sampling variability of liver biopsy in nonalcoholic fatty liver disease. *Gastroenterology* 128(7):1898–1906
- Li TQ, Aisen AM, Hindmarsh T (2004) Assessment of hepatic iron content using magnetic resonance imaging. *Acta Radiol* 45(2):119–129
- Sirlin CB, Reeder SB (2010) Magnetic resonance imaging quantification of liver iron. *Magn Reson Imaging Clin N Am* 18(3):359–381
- Bowen CV, Zhang X, Saab G, Gareau PJ, Rutt BK (2002) Application of the static dephasing regime theory to superparamagnetic iron-oxide loaded cells. *Magn Reson Med* 48(1):52–61
- Yu H, Shimakawa A, McKenzie CA, Brodsky E, Brittain JH, Reeder SB (2008) Multiecho water-fat separation and simultaneous R2\* estimation with multifrequency fat spectrum modeling. *Magn Reson Med* 60(5):1122–1134
- Clark PR, Chua-anusorn W, St. Pierre TG (2003) Bi-exponential proton transverse relaxation rate (R2) image analysis using RF field intensity-weighted spin density projection: potential for R2 measurement of iron-loaded liver. *Magn Reson Imaging* 21(5):519–530
- Bernard CP, Liney GP, Manton DJ, Turnbull LW, Langton CM (2008) Comparison of fat quantification methods: a phantom study at 30T. *J Magn Reson imaging* 27(1):192–197
- Hines CD, Yu H, Shimakawa A, McKenzie CA, Brittain JH, Reeder SB (2009) T1 independent, T2\* corrected MRI with accurate spectral modeling for quantification of fat: validation in a fat-water-SPIO phantom. *J Magn Reson Imaging JMIR* 30(5):1215–1222
- Reeder SB, Hernando D, Sharma S (2016) Phantom for iron and fat quantification magnetic resonance imaging. United States Patent
- McMullan D (1995) Scanning electron microscopy 1928–1965. *Scanning* 17(3):175–185
- Reinoso RF, Telfe BA, Rowland M (1997) Tissue water content in rats measured by desiccation. *J Pharmacol Toxicol Methods* 38(2):87–92
- Hernando D, Cook RJ, Diamond C, Reeder SB (2013) Magnetic susceptibility as a B0 field strength independent MRI biomarker of liver iron overload. *Magn Reson Med* 70(3):648–656
- Fukuzawa K, Hayashi T, Takahashi J, Yoshihara C, Tano M, Ji Kotoku, Saitoh S (2017) Evaluation of six-point modified dixon and magnetic resonance spectroscopy for fat quantification: a fat-water-iron phantom study. *Radiol Phys Technol* 10(3):349–358
- Bydder M, Hamilton G, de Rochefort L, Desai A, Heba ER, Loomba R, Schwimmer JB, Szevenyi NM, Sirlin CB (2018) Sources of systematic error in proton density fat fraction (PDFF) quantification in the liver evaluated from magnitude images with different numbers of echoes. *NMR Biomed* 31(1):e3843
- Storey P, Thompson AA, Carqueville CL, Wood JC, de Freitas RA, Rigsby CK (2007) R2\* imaging of transfusional iron burden at 3 T and comparison with 1.5 T. *J Magn Reson Imaging* 25(3):540–547
- Meloni A, Positano V, Keilberg P, De Marchi D, Pepe P, Zucarelli A, Campisi S, Romeo MA, Casini T, Bitti PP, Gerardi C, Lai ME, Piraino B, Giuffrida G, Secchi G, Midiri M, Lombardi M, Pepe A (2012) Feasibility, reproducibility, and reliability for the T\*2 iron evaluation at 3 T in comparison with 1.5 T. *Magn Reson Med* 68(2):543–551
- Yamamura J, Keller S, Grosse R, Schoennagel B, Nielsen P, Wang ZJ, Graessner J, Kooijman H, Adam G, Fischer R (2016) Iron measurements by quantitative MRI-R2\* at 3.0 and 1.5 T. In: *ISMRM*
- Doyle EK, Toy K, Valdez B, Chia JM, Coates T, Wood JC (2017) Ultra-short echo time images quantify high liver iron. *Magn Reson Med*. <https://doi.org/10.1002/mrm.26791>
- Krafft AJ, Loeffler RB, Song R, Tipirneni-Sajja A, McCarville MB, Robson MD, Hankins JS, Hillenbrand CM (2017) Quantitative ultrashort echo time imaging for assessment of massive iron overload at 1.5 and 3 Tesla. *Magn Reson Med*. <https://doi.org/10.1002/mrm.26592:n/a-n/a>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.